



SEQUENCE LISTING

<110> Genencor International, Inc.

Jones, Brian E.

Grant, William D.

Heaphy, Shaun

Grant, Susan

Rees, Helen

<120> Novel *Bacillus mHKcel* Cellulase

<130> GC797-2-PCT

<140> PCT/US2004/013257

<141> 2004-04-28

<150> US 60/467,315

<151> 2003-04-30

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 3796

<212> DNA

<213> *Bacillus* sp.

<220>

<221> misc_feature

<222> (1)...(3796)

<223> isolated from environmental water sample from
Sonachi Lake, Kenya

<221> misc_feature

<222> (1)...(3796)

<223> n = A,T,C or G

<400> 1

atcttttagag	atgagcgtga	tagtcagtgg	acctaaaaag	aactaaaaag	attttttagaa	60
gagattaaaa	cggaccgc	ccatctct	gtttattttg	atgggggatt	tgatttggag	120
acacaacgat	ctggccttgg	gtgtgtgatt	tattatgaac	aaaatgacac	gtcttaccgg	180
gtgagaaaaa	acgctaccgt	ggcgtcattg	acatcgaata	accaaggcaga	atatgccgct	240
ttgcatttag	gactaaaaag	aacttgaagg	gatcgggcgc	atcatctgcc	tatcaccatt	300
tacggtgatt	ctcaagttgt	gatcaatcag	ttaaaaggag	aatggggcgt	gtatggagga	360
ggtgttaaat	aatgggctg	accgtattt	atcaagcatt	tagctaaatt	aggcatgacc	420
gctacttata	agttaatccc	ccgtaaaagaa	aaccgcgaag	ccgatcaact	ggctacacaa	480
gcgttaaacg	ggcaagaaaat	tataagtcaa	cgtgatatac	gtgagcgtgg	tgcagattag	540
gctgcaccgc	gcaaaaaaaag	tcaacgtgtt	taggaatgga	cagggattaa	agcaacataa	600
ttctctctaa	gcaaacgttg	cgacagcaag	agagaagcat	ataaggtttt	tctgagttag	660
tctatttata	ccaatgtcca	cgtactaaat	aaacctctca	tcaaagtgg	ttttttgatt	720
aattcacttc	cactcctacc	tttatctata	taaattagtt	ccctttttgt	taataatcac	780
taattttggc	ggtatttttt	aatagaaaata	tatgctagat	tataaactag	tagccgtata	840
gaaggtggtg	attgccccta	taagagacgt	ctggcaaaca	taaaagcattc	gcattattat	900
aatcgaaagg	tggagatgag	acatgggtta	tacccaagct	aagtgtatgg	tgaaaaaaac	960

ggctttatc	tctgtttag	tgtgtcaatg	tttgttaccag	ttacatcagc	1020
tgaagatagg	gtctcttgc	cacagggtgg	tatccaatca	tatgttagcag	1080
tggctgaaat	ttaggtaata	catttgcgtgc	gataggagat	gatgaaaacag	1140
ccctcgta	acgagagaat	taatagaat	gattgcgtat	gaagggtata	1200
tatcccagtc	acatggcaaa	atcaaattggg	tggttctcca	gattatacaa	1260
ttatataatc	cggttagagc	aagtgtat	ttgggcgtt	gaggaagact	1320
gttaaatgtg	catcatgact	catggctgtg	gatgtatgtat	atggaaacata	1380
ggtgatggca	agatatacag	ctatttggg	acaattgtcg	aaaaattca	1440
ccataagttg	atgtttgaga	gtgtcaatga	gcctagggtt	acgcaggagt	1500
tcaagaaaat	catcatgctt	acttagaaga	tttaaataaag	acgttctatt	1560
agagtcagga	ggcaataatg	tgagcgccc	tttagtatttgc	cctacgatag	1620
gtctcaggat	ttactagatc	gcttgtatca	aacaatggaa	gacttggatg	1680
aattgccacg	gttcatttatt	atggcttttgc	gtccctttagt	gtcaatataa	1740
ccggtttgaa	caggagacac	aacaagatata	tatagacacg	tttgaccgtg	1800
atttacagcg	aatgggatcc	cagttgtatt	aggtgaattt	ggtttgttag	1860
aagtacggac	gtcattcagc	aaggtgagaa	attaaaattt	tttgagtttgc	1920
tctcaatgaa	cgtgatataa	cccatatgtt	atgggataac	ggtcagcatt	1980
aacttattca	ttgttatgtatc	aggaatttca	tgacatatta	aaagcgagtt	2040
ttctgctaca	gctgagtcata	atttcattca	tgtgaaggac	ggagagccaa	2100
acatatacag	ctttacttaa	acggaaatga	gctaactgcc	ctacaggcag	2160
gcttgtacta	ggagaggatt	atagagctagc	aggagacgta	ttaacgctaa	2220
cctcacaaga	ttaatttaccc	ctggccaaattt	aggaacgaat	gccccatca	2280
taattcttgg	gcagactggc	gtttcaattt	acagaatgtg	gacgtgccaa	2340
tacagatggc	tcaatatggc	attttgcgtat	ccctaccat	ttaatgggt	2400
gacgatggaa	gctgtttatg	caaaacggaga	atatgcggc	ccgcaagattt	2460
taaagaattt	ggcgaggcgt	tttccctaa	ttacgccaca	ggggaaattt	2520
agccttctt	aacgcggtac	gggatgtatga	tatccatttta	acatttcattt	2580
agagacggtg	gaatatacat	taacttttttt	tggaaattt	gtcaaggtt	2640
tgatttttaat	taatagataa	aacagcctac	ctatgcgtt	tgttggatgg	2700
tcatcttacc	aacaccgtgc	tttagaactt	tagaagtgtac	gggtttttttt	2760
gagagacaat	cctctatcaa	cagtcaccaa	tttttattca	ggaggtgtca	2820
cgttctatga	atgcataatag	tttctgacga	ataaaacatag	ttaaaaagaa	2880
ttcccgaggg	gaaggggata	atgccaacgt	attggattaa	atgacccat	2940
aaagggtttt	caagaggtgg	aatgggctc	gtttgttata	cttaatttac	3000
gtcattttgg	cgggtctact	tagtaagatg	actgacatca	taaaagagga	3060
tggcttaat	tcaattaatgc	tttaaatcacc	gagcattaat	gttgcacacc	3120
tttttattacc	ggtgggatg	aatgcggtag	attttacacc	aagtgtatgt	3180
ttactgaccc	ttttcctgtc	ctatatctt	tgcatgtgc	aactgtatgt	3240
ggctacgtct	gtccctccatt	gaacgatatg	ctgaaaaaaa	aaaattggcg	3300
caaataatgt	tatgagtgcg	tatacggata	ttgtacatgg	acatcgat	3360
tttagtaagga	gctgcctgag	tttatcaaaag	cgactttcc	tatttctcag	3420
acaccttgc	ggctggctcg	tctatgggag	gatacggggc	ttttaaattt	3480
aaccggAACG	tttcgctgc	gntgtgtcat	tatcaggtgc	agttgtatgt	3540
gtcaaccaga	ctcccttattt	gtaaacgcatt	ttgggttgg	gacgaaaatc	3600
atcttgcatt	ttttcattta	ataaaaaaat	ttggggatata	tgaaggggct	3660
tttttcaagc	gtgtgggaca	gaggactttt	tatataaaga	taatgtgaga	3720
atgcacgaca	agtgaatgcc	gatattactt	atgaagaagg	tcctgggtt	3780
cttatttggg	tagaaat				3796

<210> 2
 <211> 1716
 <212> DNA
 <213> *Bacillus* sp.

<220>
 <221> misc_feature
 <222> (1)...(1716)

<223> isolated from environmental water sample from
Sonachi Lake, Kenya

<400> 2

atgggttata	cccaagctaa	gtgtatggtg	aaaaaaacgg	tcttgtttgg	tttaattctc	60
tgttaggtg	tgtcaatgtt	tgtaccagtt	acatcagctg	aagatagggt	ctcttcgtca	120
cagggtggata	tccaaatcata	tgttagcagat	atgcaacctg	gctggaaattt	aggttaataca	180
tttgatgcga	taggagatga	tgaaacagca	tggggaaacc	ctcgtgtaac	gagagaattt	240
atagaaatga	ttgctgtatga	agggtataaa	agtattcgtt	tcccagtac	atggcaaaat	300
caaatgggtg	gttctccaga	ttatacataatt	aatgaagatt	atatcaagcg	ggttagagcaa	360
gtgatagatt	gggcgttgg	ggaagacttg	tatgtatgt	taaatgtgca	tcatgactca	420
tggctgtgga	tgtatgatat	ggaacataac	tatgtgagg	tgtatggcaag	atatacagct	480
atttgggaac	aattgtcgg	aaaattcaaa	aaccactccc	ataagttgt	gtttgagagt	540
gtcaatgagc	ctaggtttac	gcaggagtgg	ggagagattc	aagaaaatca	tcatgcttac	600
ttagaagatt	taaataagac	gttcttattat	attgtcagag	agtcaggagg	caataatgtg	660
gagcgcctt	tagtattgcc	tacgatagaa	acagccacgt	ctcaggattt	actagatcgc	720
ttgtatcaaa	caatggaaaga	cttggatgac	cctcattaa	ttgccacgg	tcattattat	780
ggctttggc	cctttatgt	caatatacgca	gggtacaccc	gttttgaaca	ggagacacaa	840
caagatatta	tagacacgtt	tgaccgtgtt	cataacacat	ttacagcgaa	tggatccca	900
gtttagttag	gtgaattttgg	tttggtaggc	tttgataaaa	gtacggacgt	cattcagcaa	960
ggtgagaaat	taaaattttt	ttagtttctc	atccatcatc	tcaatgaacg	tgatataacc	1020
catatgttat	gggataacgg	tcagcattta	aagcgagaaa	cttattcatg	gtatgtatcg	1080
gaatttcatg	acatattaaa	agcgagttgg	gagggcggtt	ctgctacagc	ttagtctaatt	1140
ttcattcatg	tgaaggacgg	agagccattt	agagatcaac	atatacagct	ttacttaaac	1200
ggaaatgagc	taactgcct	acaggcaggg	gacgaatcgc	ttgtactagg	agaggattat	1260
gagctagcag	gagacgtatt	aacgctaaaa	gccccatcc	tcacaagatt	aattaccctt	1320
ggccaattag	gaacgaatgc	ggtcatcaca	gctcaattta	attctggagc	agactggcgt	1380
tttcaattac	agaatgtgga	cgtgccaaca	gtcgaaaata	cagatggctc	aatatggcat	1440
tttgcgatcc	ctacccattt	taatgggtat	agtcttgcga	cgatggaaagc	tgtttatgca	1500
aacggagaat	atgctggccc	gcaagattgg	acgtcattta	aagaatttgg	cgaggcggtt	1560
tcccctaatt	acgccacagg	ggaaatttatt	ataacagaag	ccttctttaa	cgcggatcgg	1620
gatgatgata	tccatttaac	atttcattat	tggagcggag	agacggtgga	atatacattta	1680
cgtaaaaatg	gaaattatgt	tcaaggtaga	cggtaa			1716

<210> 3

<211> 571

<212> PRT

<213> *Bacillus* sp.

<220>

<221> VARIANT

<222> (1) ... (571)

<223> isolated from environmental water sample from
Sonachi Lake, Kenya

<400> 3

Met	Gly	Tyr	Thr	Gln	Ala	Lys	Cys	Met	Val	Lys	Thr	Val	Leu	Phe	
1				5				10			15				
Gly	Leu	Ile	Leu	Cys	Leu	Gly	Val	Ser	Met	Phe	Val	Pro	Val	Thr	Ser
								20		25		30			
Ala	Glu	Asp	Arg	Val	Ser	Ser	Gln	Val	Asp	Ile	Gln	Ser	Tyr	Val	
								35		40		45			
Ala	Asp	Met	Gln	Pro	Gly	Trp	Asn	Leu	Gly	Asn	Thr	Phe	Asp	Ala	Ile
								50		55		60			
Gly	Asp	Asp	Glu	Thr	Ala	Trp	Gly	Asn	Pro	Arg	Val	Thr	Arg	Glu	Leu
								65		70		75		80	
Ile	Glu	Met	Ile	Ala	Asp	Glu	Gly	Tyr	Lys	Ser	Ile	Arg	Ile	Pro	Val

85	90	95
Thr Trp Gln Asn Gln Met Gly	Gly Ser Pro Asp Tyr	Thr Ile Asn Glu
100	105	110
Asp Tyr Ile Lys Arg Val Glu	Gln Val Ile Asp Trp Ala	Leu Glu Glu
115	120	125
Asp Leu Tyr Val Met Leu Asn	Val His His Asp Ser Trp	Leu Trp Met
130	135	140
Tyr Asp Met Glu His Asn	Tyr Asp Glu Val Met Ala	Arg Tyr Thr Ala
145	150	155
Ile Trp Glu Gln Leu Ser	Glu Lys Phe Lys Asn His	Ser His Lys Leu
165	170	175
Met Phe Glu Ser Val Asn	Glu Pro Arg Phe Thr Gln	Glu Trp Gly Glu
180	185	190
Ile Gln Glu Asn His His	Ala Tyr Leu Glu Asp	Leu Asn Lys Thr Phe
195	200	205
Tyr Tyr Ile Val Arg Glu	Ser Gly Gly Asn Asn Val	Glu Arg Pro Leu
210	215	220
Val Leu Pro Thr Ile Glu	Thr Ala Thr Ser Gln Asp	Leu Leu Asp Arg
225	230	235
Leu Tyr Gln Thr Met Glu	Asp Leu Asp Asp Pro	His Leu Ile Ala Thr
245	250	255
Val His Tyr Tyr Gly	Phe Trp Pro Phe Ser Val	Asn Ile Ala Gly Tyr
260	265	270
Thr Arg Phe Glu Gln Glu	Thr Gln Gln Asp Ile Ile Asp	Thr Phe Asp
275	280	285
Arg Val His Asn Thr Phe	Thr Ala Asn Gly Ile	Pro Val Val Leu Gly
290	295	300
Glu Phe Gly Leu Leu	Gly Phe Asp Lys Ser	Thr Asp Val Ile Gln Gln
305	310	315
Gly Glu Lys Leu Lys	Phe Phe Glu Phe Leu Ile	His His Leu Asn Glu
325	330	335
Arg Asp Ile Thr His	Met Leu Trp Asp Asn Gly	Gln His Leu Lys Arg
340	345	350
Glu Thr Tyr Ser Trp Tyr	Asp Gln Glu Phe His Asp	Ile Leu Lys Ala
355	360	365
Ser Trp Glu Gly Arg	Ser Ala Thr Ala Glu Ser	Asn Phe Ile His Val
370	375	380
Lys Asp Gly Glu Pro	Ile Arg Asp Gln His	Ile Gln Leu Tyr Leu Asn
385	390	395
Gly Asn Glu Leu Thr	Ala Leu Gln Ala Gly	Asp Glu Ser Leu Val Leu
405	410	415
Gly Glu Asp Tyr Glu	Leu Ala Gly Asp Val	Leu Thr Leu Lys Ala Gly
420	425	430
Ile Leu Thr Arg Leu	Ile Thr Pro Gly Gln	Leu Gly Thr Asn Ala Val
435	440	445
Ile Thr Ala Gln Phe	Asn Ser Gly Ala Asp	Trp Arg Phe Gln Leu Gln
450	455	460
Asn Val Asp Val Pro	Thr Val Glu Asn Thr Asp	Gly Ser Ile Trp His
465	470	475
Phe Ala Ile Pro Thr	His Phe Asn Gly Asp	Ser Leu Ala Thr Met Glu
485	490	495
Ala Val Tyr Ala Asn	Gly Glu Tyr Ala Gly	Pro Gln Asp Trp Thr Ser
500	505	510
Phe Lys Glu Phe Gly	Glu Ala Phe Ser Pro	Asn Tyr Ala Thr Gly Glu
515	520	525
Ile Ile Ile Thr Glu	Ala Phe Phe Asn Ala	Val Arg Asp Asp Asp Ile
530	535	540

His Leu Thr Phe His Tyr Trp Ser Gly Glu Thr Val Glu Tyr Thr Leu
545 550 555 560
Arg Lys Asn Gly Asn Tyr Val Gln Gly Arg Arg
565 570